PRUTHVIRAJ MOHITE.

MBA (Information Technology and Systems Management)

Pune, India | 9096494865 | pruthvirajsantoshmohite@gmail.com |LinkedIn | Portfolio

PROFILE

Results-driven Data Scientist with 3+ years of experience, specializing in machine learning, deep learning, generative AI, data analysis, visualization, and automation. Skilled in building enterprise-grade dashboards, optimizing data workflows, and delivering actionable predictions to support business intelligence and strategic decision-making.

SKILLS

- Programming & Analytics: Python (Pandas, Scikit-learn, TensorFlow, keras), MSSQL
- Data Engineering: EDA, Data cleaning, Statistical tests, Matplotlib/Plotly, Outlier analysis, Feature engineering and feature analysis (Covariance/ Correlation, Confusion matrices)
- AI Capability: Regression/Classification Machine learning algorithms, Time Series (SARIMAX, Prophet), NLP (Text analytics using NLTK/Spacy), ANN, CNN, RNN, LSTM, Transformers (Text translation), Generative AI
- Cloud Platforms: Azure Data Factory, ML Flow, Version Control (Git) and Containerization (Docker/AWS)

EXPERIENCE

MDIndia Health Insurance TPA Pvt. Ltd. | Pune, India **Jun 2022 - Present**

Team Lead – Information & Analytics Engineering

- Developed interactive **Power BI dashboards** for operational and financial reporting, SQL Server, RLS, Modelling, DAX measures which improved reporting efficiency by 40%.
- Performed **exploratory data analysis (EDA)**, outlier detection, and data cleaning using **statistical methods** and Python libraries like **Numpy**/**Pandas**/**Matplotlib**/**Plotly**.
- Evaluated Machine learning models using standard metrics such as ROC-AUC and cross-validation.
- **NLP solutions** using **NLTK** for text preprocessing and sentiment analysis; Implemented **transformers** for advanced tasks like **text translation**.
- Developed deep learning models including **ANN, CNN, RNN, and LSTM** using **TensorFlow/Keras** for time series and sequence modeling projects.

❖ CERTIFICATIONS

- Microsoft Certified: AZ-900 Azure Fundamentals
- Microsoft Certified: DP-700 Prep. Fabric Data Engineer Associate
- Microsoft Certified: PL-300 Power BI Data Analyst
- Data Science, Machine Learning , Deep Learning and NLP Bootcamp (Udemy)

PROJECTS

- **Analytical Dashboards | Power-BI, MSSQL, DAX, M-Query**

Developed interactive Power BI dashboards to visualize policy performance, premium collections, and claims analytics across multiple brokers, enabling data-driven decision-making and performance monitoring.

Optimized complex SQL queries and stored procedures to efficiently process and transform large-scale insurance datasets, improving data refresh speed and backend performance.

Designed robust data models with calculated columns, DAX measures, and dynamic relationships to support flexible cross-filtering, drill-through analysis, and self-service BI capabilities.

Automated reporting workflows by implementing scheduled and incremental refresh in Power BI, reducing manual effort and ensuring real-time data accuracy.

- **Insurance Liability & KPI Forecasting | Python, Scikit-learn, SARIMAX, Prophet, LSTM**

Developed an **end-to-end ML pipeline** combining regression and time series models for insurance forecasting.

Implemented **feature engineering**, categorical encoding, and missing value imputation to enhance data quality and model robustness.

Trained a **regression model** ($R^2 = 0.93$) for liability amount prediction using cross-validation, grid search, and outlier impact analysis for optimal performance.

Designed a **multi-model forecasting framework** using **SARIMAX**, **Prophet**, and **LSTM** to capture trend, seasonality, and long-term dependencies in KPIs.

Conducted **ADF stationarity tests**, lag-based feature creation, and outlier correction to stabilize time series data.

- **Policy Comparison Bot**

Developed a Generative AI–based Chabot leveraging Lang Chain and OpenAI embedding's to compare and recommend insurance policies based on user-defined factors such as cost, coverage, exhaustion limits, and benefits.

Implemented document loaders and text splitters for structured ingestion and preprocessing of large policy documents, enhancing retrieval accuracy and response relevance.

Designed an intelligent QA module for follow-up queries, enabling users to ask natural language questions about policy details and receive fact-based, AI-generated answers.

Delivered a production-ready conversational system that automates policy analysis, improves decision-making, and enhances user engagement with explainable AI recommendations.